Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2015, 7(2):281-291

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Hybrid approach of Kernelized Fuzzy C-Means and Support Vector Machine for Breast Medical Image Segmentation

M. S. Sheeba[1] and A. Sathya[2]

[1]Department of Electronics & Communication Engineering, Sathyabama University, Chennai, Tamil Nadu, India
[2]Department of Mathematics, Sathybama University, Chennai, TamilNadu,India

---

ABSTRACT

Medical image analysis is becoming progressively more significant in decision making of MRI analysis and computerized diagnosis system. Image segmentation is an important key process for image analysis. In this paper, a novel medical image segmentation technique is proposed which combines Kernelized Fuzzy C-Means and Support Vector Machine. In the proposed system, a robust Hyper tangent induced Kernelized Fuzzy C-Means method is constructed with the inclusion of new spatial information term firstly. And then, the new Support Vector Machine is developed for improving effective segmentation result. The input vector for SVM classifier is generated by membership function of novel FCM in which the pixel data are labeled by new FCM method. In order to accelerate the effectiveness of segmentation result and to deal non linearity, new hyper tangent based similarity measure is used in both KFCM and SVM. Experimental analysis is carried out on real left and right breast MRIs to show the efficiency of proposed method. The performance of proposed method is demonstrated through comparative analysis of proposed and existed methods. Fuzzy Partition coefficient, Fuzzy entropy, iteration count and error rate are used to measure cluster validity. Finally, it is shown that our proposed method is the most promising technique for medical image segmentation.

Keywords: Fuzzy C-Means, Hyper tangent function, Support Vector Machine, Image Segmentation, Breast Magnetic Resonance Image, Cluster validity measure

---

INTRODUCTION

Medical imaging takes a significant part in diagnose and the treatment of patients, and final stage of drug development and validation. It is an essential implement in brain tumor and cancer detection, bone cancer detection, breast cancer detection and, etc. The remarkable developments in the field of medical imaging are blown up by the development of image processing, pattern recognition and machine intelligence techniques, and they have strengthened the recent advancement of the medical diagnostics process. There are many imaging techniques available for producing the images such as X-ray, CT scans, Positron emission tomography (PET), Ultrasound and Magnetic Resonance Imaging (MRI) to diagnosis the disease. Among them MRI has been received much attention, since it gives high contrast between different soft tissues and high spatial resolution across the entire body field of view and it has several advantages over other medical imaging techniques. However, the MRIs are extremely sensitive to physical, chemical and biological characteristics of tissues and fluids, due to the complexity of the equipments problems in imaging such as patient motion, limitation of imaging equipments, problem in receiving coils. Therefore, the ability of visualizing and detecting the abnormality of breast and bone may be affected by intensity inhomogeneities, partial volume effect, poor image contrast, high level of noises and other artifacts. This causes the necessity of image segmentation.

Image segmentation [6, 9, 27] is a process of partitioning an image into different regions such that each region is, but the union of any two adjacent regions is not, homogeneous. At the initial stage, the segmentation has been done manually by well-trained radiologists for segmenting the medical images. But the process of manual segmentation is very time-consuming, too expensive and often human errors occurred. Moreover, it does not use the complete multi-dimensional image data, and often it requires a trained expert who has wide knowledge of the anatomy being segmented to perform the segmentation process. A very important drawback of manual segmentation method is that they are labor intense, especially if there are many slices and sets of images, it will take considerable time for the user to perform the segmentation. Therefore, it needs the mathematical assisted computer based segmentation methods to segment the medical images into different tissue regions for cancer detection and efficient treatment plan in the medical analysis.

During past decades, many methods have been proposed to handle the medical image segmentation problems using unsupervised clustering methods. However, the methods have some limitations to overcome the problems of image such as intensity inhomogeneity, partial volume effect, heavy noises and other artifacts during the segmentation process. Among all other mathematical based segmentation methods, Fuzzy C-Means has received much attention since it gains the original information from image itself and it does not require any prior information to process the segmentation system.

Fuzzy C-Means is the method of fuzzy clustering technique. Fuzzy clustering is one of the most important techniques in cluster analysis. Over the years, there have been many methods and techniques developed to perform cluster analysis. Most traditional cluster analysis methods are crisp partitioning, in which every given object is strictly classified into a certain group. However, in practice, the class attributes of most objects are not strict and also ambiguous; hence it is not suitable for hard partitioning. Auspiciously, the fuzzy set theory was proposed by Lotfi. A. Zadeh [25]. It is an extension of classical set theory, and it provides a powerful tool for soft partitioning. The idea of using fuzzy set theory [5] for clustering is firstly, introduced by Ruspini [15]. Since fuzzy clustering obtains the degree of uncertainty of samples belongings to each class and expresses the intermediate property of their memberships, it can more objectively reflect the real world problems. Recently, Fuzzy clustering technique [10, 11, 14] is widely applied in many applications such as medical diagnosis, pattern recognition, data analysis and image segmentation. Since it does not require any prior information about the objects of data and any human interference of images, it is an important tool in analyzing the behavior and structural complexity of images in medical image segmentation [21, 24]. Fuzzy clustering process is carried out effectively by using FCM algorithm [12, 13].

Zhiwen Yu et al proposed a new modified support vector machine for segmenting color images in [26] to give significant reduction in the computational cost. In [3], the authors proposed self-organizing Takagi–Sugeno (T–S)-type fuzzy network with support vector (SOTFN-SV) learning based three-stage face detection method in color image segmentation. This method has given a fast detection speed and detected not only face and but also its size and orientation. An effective support vector clustering method was proposed by Jih-Jeng Huang et al [8] for market segmentation which is useful for decision making in marketing field. A Pixel wise support vector machine classification method was introduced by combining the concept of Fuzzy C-Means with Support Vector Machine (FCM based SVM) in [22] for segmenting color images. In this proposed method, the pixel level colour feature and texture future were extracted through local homogeneity model and Gabor filter and were used as input vector for SVM firstly. And then, FCM method trained SVM classifier model with extracted pixel level feature for obtaining well color image segmentation.

Juang and Hsieh [7] proposed a new classification method, the Fuzzy C-Means based support vector machine (FCM-SVM) for channel equalization. For the color image segmentation of Dam wall image, modified Fuzzy C-Means combined with Support vector regression was proposed by Dancea et al [4]. Authors [23] proposed new method, called KFCM-FSVM to correct misunderstanding of Gaussian-function-based kernel fuzzy clustering, and to deal the classification problems with outliers or noises.

This work considers the above problems and tries to get the solutions by developing effective Fuzzy C-Means (FCM) clustering methods. The new Hyper tangent kernel trick is introduced to FCM in this work for obtaining clear boundaries between the different tissues of breast MRI and thus improving the segmentation accuracy. In addition, the hyper tangent kernel based modified FCM capable of handling the general shaped dataset.

II. BASICS

A. Kernelized Fuzzy C-Means

Kernelized Fuzzy C-Means [16] is basically derived from conventional Fuzzy C-Means by replacing Euclidean distance function with kernel induced distance function. The objective function of conventional FCM [1, 2] is

$$Z(M,C) = \sum_{i=1}^{n} \sum_{k=1}^{K} m_{ik}^{f} \|p_i - c_k\|^2 \tag{1}$$

Here M represent the partition matrix $M = [m_{ik}]n \times K$, satisfies the condition

$$\sum_{k=1}^{K} m_{ik} = 1 \text{ , } for\ all\ i = 1,2,\ldots,n. \tag{2}$$

The Euclidean distance which is used in objective function (1) measures distance between the data point and cluster center. The inconvenience in using Euclidean distance is it measures only measures only noise free data and Euclidean shaped dataset. So the Euclidean distance function is replaced by kernel induced distance function to measure the distance between data point and cluster center to provide better clustering result.

The objective function (1) can be modified with the kernel induced distance as

$$Z_{kfcm}(M,C) = \sum_{i=1}^{n} \sum_{k=1}^{K} m_{ik}^{f} \|\psi(p_i) - \psi(c_k)\|^2 \tag{3}$$

For this purpose a mapping $\psi: P^d \to F$ is used whereby an object ***p*** is mapped into ***F***:

$$\psi(p) = (\psi_1(p), \psi_2(p), \ldots.) \tag{4}$$

Although ***p*** is the ***s***-dimensional vector, $\psi(p)$may have the infinite dimension. In the nonlinear classification method, an explicit form of $\psi(p)$is unavailable, but the inner product is denoted by K(p, q) and is defined as

$$K(p,q) = \langle \psi(p), \psi(q) \rangle \tag{5}$$

The function K(p,q) is known as kernel function. There are several types of Kernel function such as Gaussian, Hyper tangent, Quadratic etc. From the property of inner product

$$\|\psi(p_i) - \psi(c_k)\|^2 = \langle \psi(p_i) - \psi(c_k), \psi(p_i) - \psi(c_k) \rangle \tag{6}$$

And

$$\langle \psi(p_i), \psi(c_k) \rangle = K(p_i, c_k) \tag{7}$$

Hence we can get kernel induced distance function as

$$\|\psi(p_i) - \psi(c_k)\|^2 = K(p_i, p_i) + K(c_k, c_k) - 2K(p_i, c_k) \tag{8}$$

The kernelized Fuzzy C-means algorithm is obtained by minimizing the objective function (8) subject to the constraint (2) as processing in FCM. The equation of updated membership function and cluster center are obtained as

$$m_{ik} = \frac{1}{\sum_{j=1}^{K} \left[\frac{[1-K(p_i,c_k)]}{[1-K(p_i,c_j)]} \right]^{\frac{1}{f-1}}} \tag{9}$$

$$c_k = \frac{\sum_{i=1}^{N} m_{ik}^{f} K(p_i,c_k) y_i}{\sum_{i=1}^{N} m_{ik}^{f} K(p_i,c_k)} \tag{10}$$

B. Support Vector Machine

Support vector machine (SVM) was introduced by Vapnik, which has been used successfully in classification and problem of function estimation within the framework of statistical learning theory and structural risk minimization [18]. The convention SVM [19, 20] was constructed to separated training data into two different classes. To obtain the optimal separating hyper plane in terms of generalization error, the SVM classifier trains a training data set. The main aim of the SVM is to discover the hyper plane that maximizes the minimum distance between any data point.

Consider the training dataset of n points $\{(p_i, q_i)\}_{i=1}^{n}$ where $p_i \in R^d$ is input vector and $q_i \in \{-1, +1\}$ is the corresponding class label for the point p_i. The SVM models get form in feature space as

$$q(p) = \omega^T \psi(p) + r \tag{11}$$

Here the non linear mapping $\psi(p)$ maps the input vector into higher dimensional feature space and b denotes the bias where as ω denotes weight vector of the same dimension as the feature space. SVM model works on the following linear separable case assumptions

$$\begin{Bmatrix} \omega^T p_i + r \geq +1 \;\; if \; q_i = +1 \\ \omega^T p_i + r \leq -1 \;\; if \; q_i = -1 \end{Bmatrix} \tag{12}$$

For the non-separable case

$$\begin{Bmatrix} \omega^T \psi(p_i) + r \geq +1 \;\; if \; q_i = +1 \\ \omega^T \psi(p_i) + r \leq -1 \;\; if \; q_i = -1 \end{Bmatrix} \tag{13}$$

In this space, a linear decision surface is erected with particular properties that make sure the high generalization ability of the network. It can be possible to find out a separating hyper-plane with a maximum margin in a feature space by using non linear kernel function. It is necessary to find an existing maximum margin $\frac{2}{\|\omega\|}$ between the classes among all hyper-planes separating the data.

The problem is converted into a quadratic programming problem as

$$\min \frac{1}{2} \omega^T \omega + \gamma \sum_{i=1}^{n} \lambda_i \tag{14}$$

Such that $q_i(\omega^T \psi(p_i) + r) = 1 - \lambda_i \;\; \lambda_i \geq 0, i = 1,2, \ldots n.$ (15)

λ_i is the slack variables to tolerate misclassifications and the regularization parameter γ is a constant to trade off between the maximization of the margin and minimization of the classification error. The large γ, the more the error term is emphasized and small means that the large classification margin is encouraged.

The problem of finding the weight vector can be reformulated as minimizing the following function

$$J(\omega) = \frac{1}{2} \omega^T \omega + \gamma \sum_{i=1}^{n} \lambda_i \tag{16}$$

subject to $q_i(\omega^T p_i + r) = 1 - \lambda_i$ (17)

The quadratic problem is solved by using Lagrangian multiplier's method. The solution satisfies the Karush Kuhn-Tucker method conditions.

$$L(\omega, r, \lambda, \alpha, \beta) = \frac{1}{2} \omega^T \omega + \gamma \sum_{i=1}^{n} \lambda_i - \sum_{i=1}^{n} \alpha_i [\, q_i(\omega^T p_i + r) = 1 - \lambda_i] - \sum_{i=1}^{n} \beta_i \lambda_i \tag{18}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_n)^T$ are Lagrangian multipliers.

The solution to this quadratic programming problem is given by maximizing L with respect to $\alpha_i \geq 0$ and minimizing with respect to ω and r. Taking partial derivative with respect to ω and r, and setting the derivatives equal to zero yields

$$\frac{\partial L}{\partial \omega} = \omega - \sum_{i=1}^{n} \alpha_i \, q_i p_i = 0 \tag{19}$$

$$\frac{\partial L}{\partial r} = -\sum_{i=1}^{n} \alpha_i \, q_i = 0 \tag{20}$$

And

$$\frac{\partial L}{\partial \lambda_i} = \gamma - \alpha_i - \beta_i = 0 \tag{21}$$

So that the optimal weights are given by

$$\omega^* = \sum_{i=1}^{n} \alpha_i^* \, q_i p_i \tag{22}$$

Substituting (21) and (22) into (16) we can write

$$J(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\omega^T \omega = \sum_{i=1}^{n} \alpha_n - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \, \alpha_j q_i q_j p_i^T p_j \tag{23}$$

Note that the Lagrangian multipliers are only non-zero when $q_i(\omega^T p_i + r)$, vectors for which this is the case are called Support vectors since they lie closest to the separating hyper plane. The optimal weights are given by (22) and the bias is given by $r^* = q_i - (\omega^*)^T p_i$ (24)

Now the problem is transformed into its dual form

Maximize

$$J(\alpha) = \sum_{i=1}^{n} \alpha_n - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \, \alpha_j q_i q_j \langle p_i, p_j \rangle \tag{25}$$

subject to the constraints $\sum_{i=1}^{n} \alpha_i \, q_i = 0$ and $0 \leq \alpha_i \leq \gamma$ (26)

Then the decision function is given by

$$f(p) = sign \sum_{i=1}^{n} q_i \, \alpha_i^* \langle p, p_i \rangle + r^* \tag{27}$$

One merit of SVM is to map the input vectors into a high dimensional feature space and thus can solve the nonlinear case. In this case, the kernel trick of SVM allows us to substitute the dot product of data points above with the kernel function $Ker(p, p_i) = \langle p, p_i \rangle$. Now, the decision function is formulated as

$$f(p) = sign \sum_{i=1}^{n} q_i \, Ker(p, p_i) + r^* \tag{28}$$

f(p) indicates that the distance between the testing data p and the optimal hyper plane. Therefore, the binary classification can be solved based on the output of the *f(p)* function.

III. PROPOSED METHOD

A. Optimal Spatial neighborhood constrained Hyper tangent Fuzzy C-Means (OSHFCM)

We propose enrichment to the KFCM algorithm to improve its robustness to image segmentation. The enhancement is made by directly incorporating a new spatial constrain term and hyper tangent [17] induced distance into the objective function of KFCM that constrains the behavior of the membership function such that the membership value at each pixel depends not only on the pixel, but also on the neighborhood pixels. The effective spatial neighborhood term gives the smoothness that is depend upon the membership grade function and is determined effectively. It could be noted that the computation complexity of the proposed method is not considerably higher than that of KFCM since the added computation is only that of the spatial constraint term. The objective function of OSHFCM is as follows:

$$Z_{OSHFCM}(M,C) = 2\sum_{i=1}^{n}\sum_{k=1}^{K} m_{ik}^{f}\,[1-H(p_i,c_k)] + \frac{N}{\varepsilon}\sum_{i=1}^{n}\sum_{k=1}^{K} m_{ik}^{f}\,e^{-D_i} \tag{29}$$

where $D_i = \sum_{j=1}^{N}\|p_i - w_{ij}\|$, N is the number of neighborhood pixels and w_{ij} is the neighborhood pixels of p_i.

To get updated membership equation and cluster center equation, the objective function is minimized subject to the constraint (2) using Lagrange multipliers method. That is the first derivative with respect to membership grade and cluster center partially will set to zero to get updated estimator as follows.

a) Membership function updater

The objective function (16) is minimized using Lagrangian multiplier's method. The Lagrangian of (29) is

$$L_{OSHFCM}(M,C,\alpha) = 2\sum_{i=1}^{n}\sum_{k=1}^{K} m_{ik}^{f}\,[1-H(p_i,c_k)] + \frac{N}{\varepsilon}\sum_{i=1}^{n}\sum_{k=1}^{K} m_{ik}^{f}\,e^{-D_i} - \sum_{i=1}^{n}\alpha_i\left(\sum_{k=1}^{K} m_{ik} - 1\right) \tag{30}$$

Taking partial derivative of (30) with respect to m_{ik},

$$\frac{\partial L}{\partial m_{ik}} = 2f m_{ik}^{f-1}[1-H(p_i,c_k)] + \frac{N}{\varepsilon} f m_{ik}^{f-1} e^{-D_i} - \alpha_i = 0 \tag{31}$$

$$2f m_{ik}^{f-1}[1-H(p_i,c_k)] + \frac{N}{\varepsilon} f m_{ik}^{f-1} e^{-D_i} = \alpha_i \tag{32}$$

$$m_{ik}^{f-1} = \frac{\alpha_i}{f\left[2[1-H(p_i,c_k)] + \frac{N}{\varepsilon}e^{-D_i}\right]} \tag{33}$$

$$m_{ik} = \left[\frac{\alpha_i}{2f}\right]^{\frac{1}{f-1}}\left[\frac{1}{\left[[1-H(p_i,c_k)] + \frac{N}{2\varepsilon}e^{-D_i}\right]}\right]^{\frac{1}{f-1}} \tag{34}$$

Since $\sum_{k=1}^{K} m_{ik} = 1$,

$$\sum_{k=1}^{K}\left[\frac{\alpha_i}{2f}\right]^{\frac{1}{f-1}}\left[\frac{1}{\left[[1-H(p_i,c_k)] + \frac{N}{2\varepsilon}e^{-D_i}\right]}\right]^{\frac{1}{f-1}} = 1 \tag{35}$$

$$\left[\frac{\alpha_i}{2f}\right]^{\frac{1}{f-1}} = \frac{1}{\sum_{k=1}^{K}\left[\frac{1}{\left[[1-H(p_i,c_k)] + \frac{N}{2\varepsilon}e^{-D_i}\right]}\right]^{\frac{1}{f-1}}} \tag{36}$$

Substituting (36) in (34), we get

$$m_{ik} = \frac{1}{\sum_{j=1}^{K}\left[\frac{\left[[1-H(p_i,c_k)] + \frac{N}{2\varepsilon}e^{-D_i}\right]}{\left[[1-H(p_i,c_j)] + \frac{N}{2\varepsilon}e^{-D_j}\right]}\right]^{\frac{1}{f-1}}} \tag{37}$$

This equation can provide effective membership grade in successive iteration.

b) Cluster center updater

To get cluster center updater equation, differentiating the objective function (29) partially with respect to c_k and equating to zero as follows: (29) can be written as

$$Z_{OSHFCM}(M,C) = 2\sum_{i=1}^{n}\sum_{k=1}^{K} m_{ik}^{f}\tanh\left(-\frac{\|p_i-c_k\|^2}{\delta^2}\right) + \frac{N}{\varepsilon}\sum_{i=1}^{n}\sum_{k=1}^{K} m_{ik}^{f} e^{-D_i} \tag{38}$$

$$\frac{\partial J}{\partial c_k} = 2\sum_{i=1}^{n} sech^2\left(-\frac{\|p_i-c_k\|^2}{\delta^2}\right)\left(-2\frac{p_i-c_k}{\delta^2}\right)(2) + \frac{N}{\varepsilon}\sum_{i=1}^{N} m_{ik}^f\, e^{-D_i} = 0 \quad (39)$$

$$\sum_{i=1}^{N} m_{ik}^f\left[1-\tanh(-\frac{\|p_i-c_k\|^2}{\delta^2})\right]\left[1+\tanh(-\frac{\|p_i-c_k\|^2}{\delta^2})\right]\left(\frac{p_i+c_k}{\delta^2}\right) + \frac{N}{4\varepsilon}\sum_{i=1}^{N} m_{ik}^f\, e^{-D_i} = 0 \quad (40)$$

Simplifying this equation, we will get updating cluster center equation as

$$c_k^{new} = \frac{\sum_{i=1}^{N} m_{ik}^f\left[H\left(p_i,c_k^{old}\right)T\left(p_i,c_k^{old}\right)\frac{1}{\delta^2}+\frac{N}{\varepsilon}e^{-D_i}\right]p_i}{\sum_{i=1}^{N} m_{ik}^f\left[H(p_i,c_k^{old})T(p_i,c_k^{old})\frac{1}{\delta^2}+\frac{N}{\varepsilon}e^{-D_i}\right]} \quad (41)$$

where $T(p_i,c_k) = 1+\tanh(-\frac{\|p_i-c_k\|^2}{\delta^2})$

B. Proposed Support Vector Machine

The proposed KFCM provides input vector for SVM process when a number of cluster center and initial cluster centers are given. Consider the membership function matrix as $M = [m_k(p_i)]$ for the training data $\{p_1, p_2, ..., p_n\}$.

When a number of cluster K and initial cluster center are given, the proposed FCM algorithm is employed for cluster construction. It provides the label for data and membership function. Let the set of training data be $\{p_1, p_2, ..., p_n\}$. Through the membership function (37), each input data p_n is transformed to the vector $m(p_i) = [m_1(p_i), m_2(p_i), \ldots, m_k(p_n)]$ where $m_k(p_n)$ is the output of the n^{th} data point. The vector m is the input to the SVM, and the training data points are represented by

$$T = \{(m(p_1),q_1),(m(p_2),q_2),\ldots,(m(p_n),q_n)\} \quad (42)$$

Then the decision function of SVM classifier is written as

$$F(p) = sign\sum_{i=1}^{n} q_i\,\alpha_i^*\langle m(p), m(p_i)\rangle + r^* \quad (43)$$

where α_i^* is solved from (27)

$\langle m(p), m(p_i)\rangle$ can be replaced by hyper tangent kernel function using the relation between of kernel function and inner product. (i.e) $\langle \psi(p_i), \psi(c_k)\rangle = K(p_i,c_k)$. Here we use hyper tangent function as kernel function.

Therefore, the output function in (43) can be represented as

$$F(p) = sign\sum_{i=1}^{n} q_i\,\alpha_i^*\tanh(-\frac{\|p-p_i\|^2}{\delta^2}) + r^* \quad (44)$$

The weighting parameters of proposed FCM-SVM are calculated by (22) and (24).

IV. EXPERIMENTAL STUDY

A. Experimental work on real breast images

In order to prove the effectiveness of the proposed method, the real left and right breast MRIs are considered in this section. The existed methods such as KFCM, FCM based SVM, FCM-SVM, KFCM-FSVM and proposed method, are executed on the images under the same initial conditions. For the experimental purpose, the Gaussian noise is inserted in real breast MRI. The Gaussian noise corrupted images are given in Fig. 1(a) & 2(a). Using the algorithms the images are segmented into four tissues such as fat, normal tissue, benign lesions and malignant lesions. The various colors are used to identify different tissues that are grey for fat tissue, blue for normal tissue, red for benign lesions and green for malignant lesions.

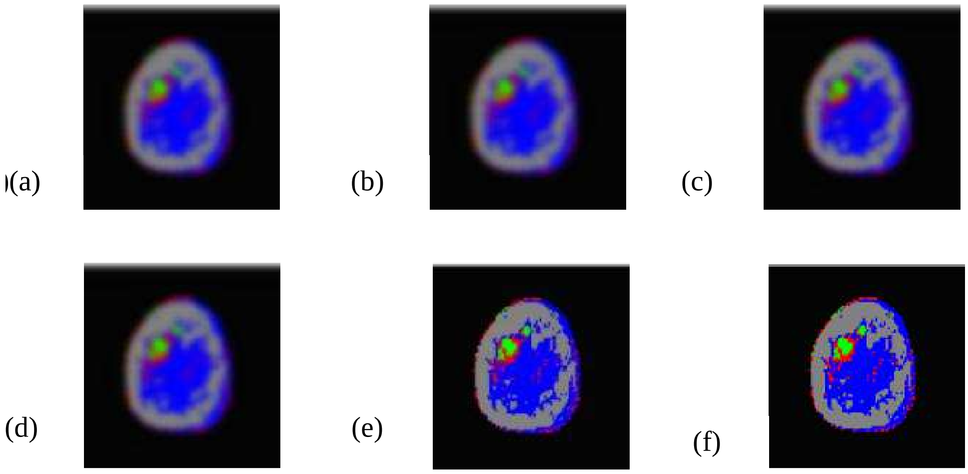

Figure 1: (a) Left breast MRI corrupted by Gaussian noise (b) Segmented Image by KFCM (c) Segmented Image by FCM based SVM (d) Segmented Image by FCM-SVM (e) Segmented Image by KFCM -FSVM (f) Segmented Image by proposed method

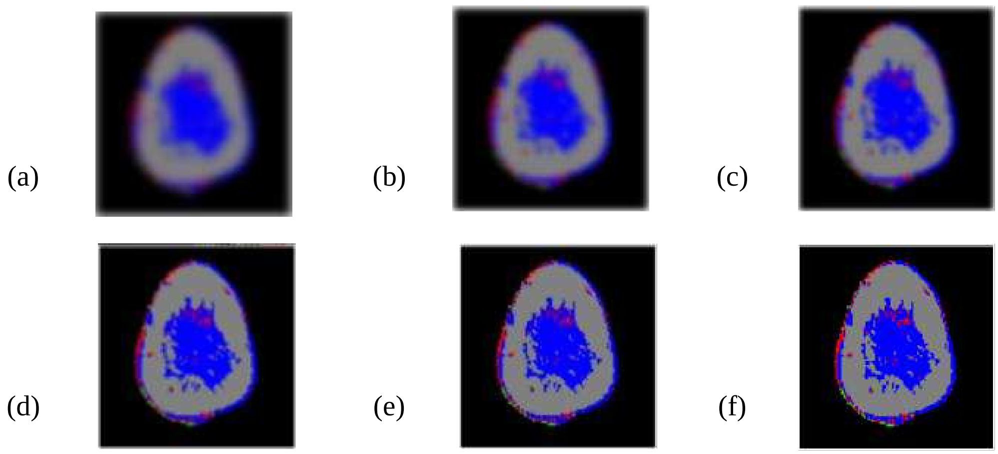

Figure 2: (a) Right breast MRI corrupted by Gaussian noise (b) Segmented Image by KFCM (c) Segmented Image by FCM based SVM (d) Segmented Image by FCM-SVM (e) Segmented Image by KFCM -FSVM (f) Segmented Image by proposed method

Figs. 1(b-f) &2(b-f) represent the segmentation results of four existed methods that are KFCM, FCM based SVM, FCM-SVM, KFCM-FSVM and proposed method executed on Gaussian noise corrupted real left and right breast MRIs. We add a Gaussian noise just for proving the effectiveness and for comparative analysis but in nature the MRIs are not affected by such a noise. Figs. 1(b-d) &2 (b-d) shows that conventional KFCM has given a poor result in the presence of Gaussian noise as well as FCM based SVM and FCM-SVM have also provided almost same result of KFCM. It can be noted that KFCM-FSVM tried to remove noise in the MRIs but failed to remove the noise completely which is given in Fig. 1(e) & 2(e). Because of novel spatial constraint term and effective distance function, our proposed method succeeds well in removing noise and provides better segmentation result than other existed methods in fig.1(f) & 2(f).

On the whole, from the Figs. 1(b-f) and 2(b-f) we can easily understand that the existed methods are very sensitive to noises where as our proposed method still more stable and attain the best segmentation performance among the methods.

B. Quantitative analysis of Segmentation result

In order to measure the segmentation accuracy, we use Fuzzy Partition coefficient and Fuzzy entropy methods in this section. Fuzzy Partition coefficient is defined as follows:

$$F_{pc} = \frac{\sum_{i=1}^{n}\sum_{k=1}^{K} m_{ik}^2}{n} \quad (45)$$

And, Fuzzy entropy is defined as $F_{pe} = \frac{-\sum_{i=1}^{n}\sum_{k=1}^{K} m_{ik} log m_{ik}}{n}$ (46)

The above partition coefficient validity method measures the fuzziness of the partition without considering dataset itself. A good partition is realized by the maximum value of fuzzy partition coefficient in the significance of a least fuzzy clustering. Fuzzy partition entropy measure gives the information about the membership matrix not considering the dataset. In this method, the minimum value represents good partition in the meaning of a more crisp partition.

So it can be realized that the algorithm provides best segmentation result when *F*pc attains maximum value or *F*pe attains minimum value.

Table 1 gives the values of *F*pc and *F*pe of the proposed method and four existed method on real breast MRI and knee bone MRI. Further, Table 2 gives the iteration count and error rate of proposed method and the existed methods, where the error rate can be calculated based number of misclassified objects and the total number of objects in the data. That is

$$Error\ (in\ \%) = \frac{Number\ of\ misclassified\ patterns}{Total\ number\ of\ patterns} \times 100 \quad (47)$$

The misclassified objects are identified from the following formula:

$$M(p_i) = \frac{B_i - B_i^{'}}{\max\{B_i, B_i^{'}\}} \quad (48)$$

Here,

$$B_i = min_k\left\{A_i^{(k)} / \ i = 1,2,\dots,N, k = 1,2,\dots,K\right\} where\ A_i^{(k)} = Average\{Dist(y_i, c_k)\}$$

$Dist(p_i, c_k)$ is the distance between y_i and all the elements of c_k. $B_i^{'}$ = Average of distance measure between p_i and all other elements in the same cluster. In this $M(p_i) \in [-1,1]$. If $M(p_i)$ is close to -1, then the element p_i is considered as misclassified objects.

Table 1: Comparison of the segmentation result on real motion artifact breast MRI

Name of the Methods	Real right breast MRI F_{pc}	Real right breast MRI F_{pe}	Real left breast MRI F_{pc}	Real left breast MRI F_{pe}
KFCM	0.6923	0.5132	0.7029	0.5129
FCM based SVM	0.7234	0.4846	0.7192	0.4956
FCM-SVM	0.7982	0.3523	0.8103	0.3413
KFCM-FSVM	0.8723	0.2956	0.8815	0.2547
Proposed method	0.9752	0.1167	0.9812	0.1025

Table 2: Iteration count and error rate

Name of the Methods	Real right breast MRI Iteration count	Real right breast MRI Error rate	Real left breast MRI Iteration count	Real left breast MRI Error rate
KFCM	65	35%	64	35%
FCM based SVM	58	29%	59	27%
FCM-SVM	42	21%	39	20%
KFCM-FSVM	30	18%	28	17%
Proposed method	16	7%	13	5%

From the table 1, we can see that the existed methods had lower value for fuzzy partition coefficient and higher value for fuzzy entropy which represents the poor segmentation accuracy. Further, the efficiency of our proposed method has been proved through obtaining highest fuzzy partition coefficient value and lowest fuzzy entropy value which shows the best segmentation accuracy.

The iteration count and error rate of existed methods and proposed method are given table 2. In this we can realize that our proposed method has best convergence speed to run the algorithm that any other methods. Also, the proposed method has best segmentation accuracy in segmenting real left and right breast MRIs.

This experimental study clearly proved that our proposed method is promising technique for segmenting medical images.

CONCLUSION

This work analyzed the problem of segmenting medical images using Fuzzy C-Means based SVM method. In this work we proposed a new framework of hyper tangent based Fuzzy C-means and novel Support Vector Machine to improve the segmentation accuracy and to deal ambiguity in the segmentation. A new spatial constraint term was included in the objective function of hyper tangent based FCM in order remove the noises in the breast MRI. To deal the non-linearity, hyper tangent induced distance function was used in both FCM and SVM. The proposed method and existed methods were applied on real right and left breast MRI and the results were compared both visually and quantitatively to prove the effectiveness of proposed method. This paper suggests that our proposed framework is promising technique for medical image segmentation.

REFERENCES

[1] J.C Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, (**1981**) New York, Plenum Press.
[2] J.C Bezdek and S.K. Pal, *Fuzzy Models for Pattern Recognition*, (**1992**) New York: IEEE Press.
[3] Chia-Feng Juang and Shen-Jie Shiu, *Neurocomputing* 71 (**2008**) 3409–3420.
[4] O. Dancea et al., A. *IEEE International Conference on Automation Quality and Testing Robotics (AQTR), (**2010**), vol. 3, pp.1-6.*
[5] J.C Dunn, *Journal* of Cybernetics (**1973**) 3 pp.32-57.
[6] Feng Zhao, Licheng Jiao, Hanqiang Liu and Xinbo Gao, *Signal Processing*, (**2011**) No. 91, pp.988–999.
[7] C.F. Juang and C.D Hsieh, *International Journal of General Systems*, Volume 38, Issue 3 April **2009** , pages 273 – 289.
[8] Jih-Jeng Huang, Gwo-Hshiung Tzeng and Chorng-Shyong Ong, *Expert Systems with Applications* 32 (**2007**) 313–317.
[9] S.R. Kannan, A. Sathya, S. Ramathilagam, *Journal of Systems and Software*, (**2010**) 83, 2487–2495.
[10] S.R Kannan, A. Sathya, S. Ramathilagam, *International Journal of Soft Computing* (Springer Publication), (**2009**) Vol.8(4) pp. 1599–1606.
[11] S. R. Kannan et al *Computer in Biology and Medical*, (**2013**) 43(2): 73-83.
[12] S. R. Kannan, et al: *Comput. J. (***2013***), 56*(3): 393-406.
[13] Karan Sikka, Nitesh Sinha, Pankaj K. Singh and Amit K. Mishra, *Magnetic Resonance Imaging* (**2009**)27, 994–1004.
[14] S. Ramathilagam et al., *Journal of Intelligent and Fuzzy Systems*, (**2014**) 27(5): 2573-2595.
[15] Ruspini E.H, *Information and Control*, (**1969**)15(1) pp.22-32.
[16] A Sathya et al, *IEEE Digital Library*, (**2012**), pp.62-72.
[17] A Sathya, Anudevi Samuel, and M.S. Sheeba, Robust Fuzzy C-Means based Minimal Spanning tree method For Segmentation of Breast MRI, International Conference on Mathematical Sciences, Elsevier Publications, (**2014**) pp. 495-501.
[18] B.Scholkopf and A.J.Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, MIT Press, Cambridge, MA, **2002**.
[19] V.N. Vapnik, *Statistical Learning Theory, (***1998**) New York, NY: JohnWiley & Sons.
[20] V.N. Vapnik, *The Nature of Statistical Learning Theory, (***2000**) 2nd ed. New York, NY: Springer.
[21] Weijie Chen, L. Giger Maryellen and Ulrich Bick, *Academic Radiology* (**2006**)13(1), 63-72.
[22] Xiang-Yang Wang, Ting Wang and Juan Bu, *Pattern Recognition* 44 (**2011**) 777–787.
[23] Xiaowei Yang., et al, *IEEE TRANSACTION ON FUZZY SYSTEMS*, (**2011**) Vol.19 , Issue: 1, pp. 105 – 115.
[24] Ye Xing et al., Simultaneous Estimation and Segmentation of T1 Map for Breast parenchyma Measurement, 4th *IEEE International Symposium on Biomedical Imaging*, (**2007**),pp. 332 – 335.

[25] Zadeh L.A, Fuzzy sets. Inf. Control (8), (**1965**), pp. 338–353.
[26] Zhiwen Yu, Hau-San Wong and Guihua Wen, A modified support vector machine and its application to image segmentation, *Image and Vision Computing*, (**2011**), vol. 29, pp.29–40.
[27] Zhi MinWang, YengChaiSoh, QingSong, KangSim, *Pattern Recognition* (**2009**), 42, 2029 – 2044.